Computational lens for the near field

Carney, P. Scott; Franzin, Richard A.; Bozhevolnyi, Sergey I.; Volkov, Valentyn S.; Boltasseva, Alexandra; Schotland, John C.

Published in:
Physical Review Letters

Link to article, DOI:
10.1103/PhysRevLett.92.163903

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

Computational Lens for the Near Field

P. Scott Carney and Richard A. Frazin
Department of Electrical and Computer Engineering and The Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

Sergey I. Bozhevolnyi and Valentyn S. Volkov
Department of Physics and Nanotechnology, University of Aalborg, Aalborg Øst, Denmark

Alexandra Boltasseva
Research Center COM, Technical University of Denmark, Denmark

John C. Schotland
Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania, USA
(Received 22 November 2003; published 23 April 2004)

A method is presented to reconstruct the structure of a scattering object from data acquired with a photon scanning tunneling microscope . The data may be understood to form a Gabor type near-field hologram and are obtained at a distance from the sample where the field is defocused and normally uninterpretable. Object structure is obtained by the solution of the inverse scattering problem within the accuracy of a perturbative, two-dimensional model of the object.

DOI: 10.1103/PhysRevLett.92.163903 PACS numbers: 42.30.Wb, 42.25.Fx

Near-field optical microscopy has developed dramatically in recent years [1–3]. The seminal idea dates to 1928, when Synge [4] proposed that a subwavelength resolved image could be obtained by scanning a subwavelength aperture over a thin sample. Today this method is known as near-field scanning optical microscopy (NSOM) [5–9]and is practiced in many variations. In a related technique, photon scanning tunneling microscopy (PSTM), the object is illuminated by an evanescent wave generated at the face of a prism or slide, and the field is detected via a fiber probe in the near zone of the sample (as in NSOM).

In all scanning-probe modalities, the probe must be kept in near contact with the sample in order to obtain images in which structure may be recognized. Even then, the connection between the measured signal and the sample properties has proven to be problematic. When the probe is withdrawn a significant fraction of a wavelength, the images become blurred [8], a result of the fact that the field being measured is not focused. To clarify the meaning of the measurements, it is desirable to find a solution to the inverse scattering problem (ISP). Results in this direction have been reported for the case of surface profile reconstruction in homogeneous media [10–13]. There has also been recent progress for the case of three-dimensional inhomogeneous media [14–16]. In these references it is demonstrated that though resolution necessarily degrades with distance from the measurement plane (due to the loss of the evanescent modes), subwavelength tomographs may be computed from data obtained at distances up to roughly a wavelength away from the sample depending on the level of noise.

Solution of the ISP requires holographic measurement of the scattered field. PSTM measurements are inherently holographic [17]; the sample is illuminated by an evanescent field and that same field serves as a reference wave. The measurements are thus the near-field equivalent of a Gabor hologram [18]. The hologram may be measured a large fraction of a wavelength from the sample, where the near field is normally out of focus, as may be seen in Ref. [8]. It is demonstrated here that the data thus obtained, though not amenable to direct interpretation, provide enough information to determine quantitatively the two-dimensional structure of a thin sample, thus achieving a computational lens for the near field.

The PSTM experiment may be modeled by considering a field in two half-spaces with a scattering object, the sample, in one half-space. One half-space, $z \geq 0$, has the vacuum index of refraction while the $z < 0$ half-space has an index of refraction n. The sample is described by a susceptibility $\eta(\mathbf{r})$ which is nonzero only in the $z \geq 0$ region. A monochromatic evanescent plane wave generated by total internal reflection in the $z < 0$ region is incident on the sample. The resulting total field intensity is measured on a plane $z = z_d$ passing through the near zone of the sample. The sample considered here is nonmagnetic and so it is sufficient to treat only the electric field $\mathbf{E}$. The field is taken to consist of two parts,

$\mathbf{E} = \mathbf{E}^i + \mathbf{E}^s$, where $\mathbf{E}^i$ is the incident field and $\mathbf{E}^s$ is the scattered field. The incident field $\mathbf{E}^i$ obeys the reduced wave equation with $\eta = 0$. The scattered field satisfies the integral equation

$$E^s_\alpha(\mathbf{r}) = k_0^2 \int d^3r'\, G_{\alpha\beta}(\mathbf{r}, \mathbf{r}') E_\beta(\mathbf{r}') \eta(\mathbf{r}'), \qquad (1)$$

where the summation convention over repeated indices applies and will throughout. The Green's tensor $G_{\alpha\beta}(\mathbf{r}, \mathbf{r}')$ may be expressed in a plane wave decomposition of the form [3]

$$G_{\alpha\beta}(\mathbf{r}, \mathbf{r}') = \frac{i}{2\pi} \int \frac{d^2q}{k_z(\mathbf{q})} g_{\alpha\beta}(\mathbf{q}, z') e^{i\mathbf{k}(\mathbf{q})\cdot(\mathbf{r}-\mathbf{r}')}, \qquad (2)$$

where $\mathbf{k}(\mathbf{q}) = (\mathbf{q}, k_z(\mathbf{q}))$ and $k_z(\mathbf{q}) = \sqrt{k_0^2 - q^2}$. Expressions for $g_{\alpha\beta}$ may be found in [3,16]. The plane wave modes appearing in Eq. (2) are labeled by the transverse part of the wave vector, $\mathbf{q}$. The modes for which $|\mathbf{q}| < k_0$ are homogeneous, or propagating, plane waves. When $|\mathbf{q}| > k_0$, the plane wave is evanescent, decaying exponentially with increasing values of z. These waves are superoscillatory in the transverse plane and thus provide a means to probe the high spatial frequency structure of the sample.

The scattered field is computed perturbatively. In the $z \geq 0$ region, the incident field is an evanescent wave with polarization $\mathbf{e}$ so that $E^i(\mathbf{r}) = \mathbf{e} \exp[i\mathbf{k}(\mathbf{q}_i) \cdot \mathbf{r}]$, where $\mathbf{q}_i$ is the tranverse wave vector of the incident wave. It is assumed that the probe tip acts as a pointlike scatterer of susceptibility $\chi\delta(\mathbf{r} - \mathbf{r}_p)$ and that the sample is weakly scattering. There are thus two significant terms in the expression for the scattered field at the location $\mathbf{r}_p$ of the probe

$$E^s_\alpha(\mathbf{r}_p) = k_0^2 \int d^3r'\, G_{\alpha\beta}(\mathbf{r}_p, \mathbf{r}') e_\beta e^{i\mathbf{k}(\mathbf{q}_i)\cdot\mathbf{r}'} \eta(\mathbf{r}') + \chi k_0^4 \int d^3r'\, G_{\alpha\beta}(\mathbf{r}_p, \mathbf{r}') \times G_{\beta\gamma}(\mathbf{r}', \mathbf{r}_p) e_\gamma e^{i\mathbf{k}(\mathbf{q}_i)\cdot\mathbf{r}_p} \eta(\mathbf{r}'). \qquad (3)$$

The first term is the usual expression for the scattered field computed within the accuracy of the first Born approximation. The second term arises as a result of the interaction of the incident field first with the probe, then subsequently with the sample.

The total intensity, $I = |\mathbf{E}|^2$, is measured in the plane $z = z_d$. The intensity is the sum of four terms,

$$I = |\mathbf{E}^i|^2 + |\mathbf{E}^s|^2 + \mathbf{E}^s \cdot \mathbf{E}^{i\star} + \mathbf{E}^i \cdot \mathbf{E}^{s\star}, \qquad (4)$$

where the spatial arguments of the fields have been supressed. The first term is constant. The second term is quadratic in the susceptibility of the sample and thus is negligible compared to the third and fourth terms for values of z_d that are less than about a wavelength. The third and fourth terms are complex conjugates that carry the image information.

The sample is assumed to be of constant thickness Δz and to depend only on the transverse spatial coordinate so that $\eta(\mathbf{r})$ may be expressed as $\eta(\boldsymbol{\rho})$ for $0 \leq z < \Delta z$ and where $\boldsymbol{\rho}$ indicates the projection of $\mathbf{r} = (\boldsymbol{\rho}, z_d)$ in the constant z plane. By taking the Fourier transform, $\tilde{I}(\mathbf{q}, z_d)$, of $I(\boldsymbol{\rho}, z)$, the data are simply related to the Fourier transform of η. Including only the third and fourth terms of Eq. (4), it yields

$$\tilde{I}(\mathbf{q}) = K(\mathbf{q})\tilde{\eta}(\mathbf{q}) + K^\star(-\mathbf{q})\tilde{\eta}^\star(-\mathbf{q}), \qquad (5)$$

where $\tilde{\eta}(\mathbf{q})$ is the inverse Fourier transform of η with argument $\mathbf{q}$ and the z_d dependence is implicit, and K is given by the expression

$$K(\mathbf{q}, z_d) = ik_0^2 e_\alpha e_\beta \Big\{ 2\pi \frac{e^{i[k_z(\mathbf{q}_i - \mathbf{q}) - k_z^\star(\mathbf{q}_i)]z_d}}{k_z(\mathbf{q}_i - \mathbf{q})} f_{\alpha\beta}[\mathbf{q}_i - \mathbf{q}, k_z(\mathbf{q}_i) - k_z(\mathbf{q}_i - \mathbf{q})] - \chi k_0^2 e^{-2\mathrm{Im}[k_z(\mathbf{q}^i)]} \times \int \mathrm{d}^2q\, \frac{e^{\mathrm{i}[k_z(\mathbf{q}) + k_z(\mathbf{q}+\mathbf{Q})](z_d)}}{k_z(\mathbf{q}) k_z(\mathbf{q}+\mathbf{Q})} \times f_{\alpha\gamma}[\mathbf{q}, -k_z(\mathbf{q}) - k_z(\mathbf{q}+\mathbf{Q})] g_{\gamma\beta}(\mathbf{Q}+\mathbf{q}, z_d) \Big\}, \qquad (6)$$

where both $f(\mathbf{q}, p)$ and $g(\mathbf{q}, p)$ are of the form

$$M(\mathbf{q}, p) = \frac{1}{|\mathbf{q}|^2} \begin{pmatrix} q_x^2 m_{xx} + q_y^2 m_{yy} & q_x q_y (m_{xx} - m_{yy}) & |\mathbf{q}| q_x m_{xz} \\ q_x q_y (m_{xx} - m_{yy}) & q_y^2 m_{xx} + q_x^2 m_{yy} & |\mathbf{q}| q_y m_{xz} \\ |\mathbf{q}| q_x m_{zx} & |\mathbf{q}| q_y m_{zx} & |\mathbf{q}|^2 m_{zz} \end{pmatrix}, \qquad (7)$$

where

$$m(\mathbf{q}, p) = \frac{1}{k_0^2} \begin{pmatrix} k_z^2(\mathbf{q})[\zeta_1 + R_2\zeta_2] & 0 & -|\mathbf{q}| k_z(\mathbf{q})[\zeta_1 - R_2\zeta_2] \\ 0 & k_0^2[\zeta_1 + R_1\zeta_2] & 0 \\ -|\mathbf{q}| k_z(\mathbf{q})[\zeta_1 + R_2\zeta_2] & 0 & |\mathbf{q}|^2[\zeta_1 - R_2\zeta_2] \end{pmatrix}, \qquad (8)$$

and where the tensor f is obtained by taking $\zeta_1 = -\mathrm{i}e^{\mathrm{i}p\Delta z} - 1/p$, $\zeta_2 = -\mathrm{i}e^{\mathrm{i}[p + 2k_z(\mathbf{q}_i)]\Delta z} - 1/p + 2k_z(\mathbf{q}_i)$, and the tensor g is obtained by taking $\zeta_1 = 1$, and $\zeta_2 = e^{2\mathrm{i}k_z(\mathbf{q})p}$. The quantities $R_1(\mathbf{q})$ and $R_2(\mathbf{q})$ are the reflection coefficients given by $R_1(\mathbf{q}) = [k_z(\mathbf{q}) - k_z'(\mathbf{q})]/[k_z(\mathbf{q}) + k_z'(\mathbf{q})]$, and $R_2(\mathbf{q}) = [k_z'(\mathbf{q}) - nk_z(\mathbf{q})]/[k_z'(\mathbf{q}) + nk_z(\mathbf{q})]$ with $k_z'(\mathbf{q}) = \sqrt{n^2k_0^2 - q^2}$.

When only one data set is available, the problem of solving for $\tilde{\boldsymbol{\eta}}$ is underdetermined. This situation is analogous to the conjugate image problem presented in far-field holography [19]. The ambiguity may be resolved with a second data set obtained with a new illuminating field (distinct $\mathbf{q}_i$). The resulting system of equations may be expressed as $\tilde{\mathbf{I}} = \hat{K}\,\tilde{\boldsymbol{\eta}}$, where $\hat{K}$ is a matrix whose entries are given by Eq. (6). The inverse may be expressed

$$\tilde{\boldsymbol{\eta}} = (\hat{K}^{\dagger}\hat{K})^{-1}\hat{K}^{\dagger}\tilde{\mathbf{I}}, \tag{9}$$

where the $\dagger$ indicates Hermitian adjoint. To obtain a stable solution, it is necessary to regularize the solution by some means such as the Tikhonov method or, as is done here, by imposing a cutoff on the singular values used to construct $(\hat{K}^{\dagger}\hat{K})^{-1}$.

The experimental apparatus (see Fig. 1) consists of a collection NSOM with an uncoated, sharp fiber tip used as a probe [20] and an arrangement for illumination of a sample by either of two counter-propagating evanescent waves, a slight modification of the setup described in detail in Ref. [21]. A beam from a 10 mW He-Ne laser (with $\lambda = 633$ nm) is split into two beams that are coupled into a prism such that counterpropagating evanescent waves of equal magnitude are established at the prism face. The magnitude of the transverse wave vector of each illuminating evanescent wave is $|\mathbf{q}_i| = 1.05k_0$. Both beams are TE-polarized; i.e., the electric fields are perpendicular to the plane of incidence. The prism base supports the sample with an index-matching oil facilitating the optical contact between the prism and the sample. The sample is then illuminated by one beam at a time with the other beam blocked.

The field intensity is measured with the NSOM fiber probe, which can be scanned along the sample surface at a constant distance along a plane approximately parallel to the surface plane (within 15.0 nm). The probed optical field components are weakly coupled into the fiber and detected with a photomultiplier tube. The signal is proportional to the near-field intensity distribution only if the NSOM tip measures with the same efficiency all spatial frequency components of the total optical field [21]. The detection efficiency for uncoated fiber tips decreases with the increase of spatial frequency [22], resulting in an effective center of detection located inside the tip at a certain distance from the tip end [23]. This model partly accounts for the finite size of the tip and has given good agreement between experimental data and theoretical simulations with the center-tip distance being found in the range from 100 [23] up to 500 nm [24].

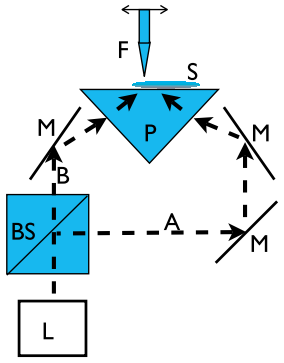

FIG. 1 (color online). The PSTM instrument. A He-Ne laser (L) emits a collimated beam that is divided at the beam splitter (BS) and propagated on one of two possible paths, (A) or (B), the other path being blocked. The beam is routed by mirrors (M) to the prism (P) at an angle such that the beam in the prism is totally internally reflected. The evanescent field outside the prism interacts with the sample (S) and the total field is detected by the tapered fiber probe (F), which is scanned in a plane at fixed height.

The sample used in this work was prepared by coating a glass substrate with a resist layer and using electron beam lithography for nanopatterning. The exposed and developed resist layer was covered with an ~80 nm-thick gold film, and the subsequent liftoff resulted in submicron-sized gold islands located on the glass substrate surface. One of the fabricated islands from the area was then selected for NSOM imaging.

Data were collected with an actual tip height 100 nm ($0.16\lambda_0$). The collected data are shown in Figs. 2(a) and 2(b). Without further processing of these data, it is difficult to make a meaningful assessment of the physical structure of the sample. Some features in each data set may be associated with features seen in the atomic force microscope (AFM), but there is little agreement between the data sets. Furthermore, the data, in normalized photocounts, do not immediately reveal the quantitative optical properties of the sample.

The object structure was computed from the data using Eq. (9). The reconstructions were computed with an effective tip offset of 200 nm; i.e., the effective height was $z_d = 300$ nm. The sample was assumed to be of uniform thickness of 80 nm. The algorithm was regularized by imposing a cutoff for singular values of the forward problem below 10% of the largest singular value to obtain a stable reconstruction in the presence of noise. The coupling parameter χ was chosen to obtain optimal reconstructions at a value of $\chi = 2.5 \times 10^7$ nm^3. After precomputation of the inverse scattering kernel, images were computed on a modern desktop computer in a few minutes. The real part of the reconstructed susceptibility is shown in Fig. 2(c). Figure 2(d) shows an AFM image of the gold deposited on a glass slide. The results are in quantitative agreement with published values of the real susceptibility of gold Re $\eta = -0.94$ at 633 nm. The reconstructed value of the real susceptibility varies in the central peak between -0.8 and -1.0.

Reconstruction resolves the ambiguity in the interpretation of two disparate data sets. The computed image shows good morphological agreement with the AFM image. While select features of the original data sets

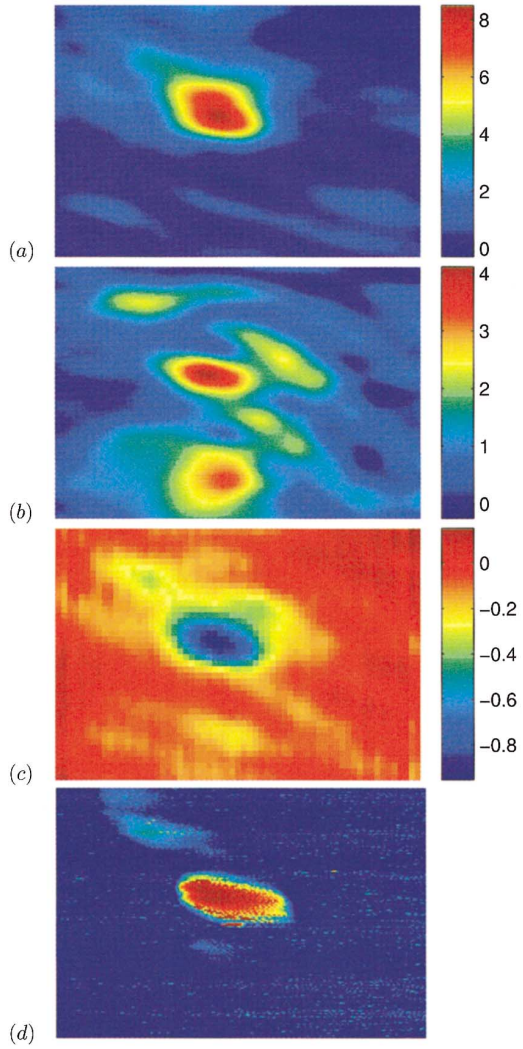

FIG. 2 (color). The data are displayed in panels (a) and (b). The colorbar indicates normalized photocount with the constant background subtracted. The result of the reconstruction algorithm is displayed in panel (c). An AFM image is displayed in panel (d) for comparison. The field of view in each image is $3.4 \times 5\ \mu m^2$.

resemble certain features seen in the AFM image, the reconstruction more closely matches the structure seen in the AFM image. The reconstruction also more clearly displays the small feature above and to the left of the main, central feature.

Computational methods can greatly enhance and extend the imaging capability of PSTM. Computed reconstruction of the object structure obviates the need to be in direct contact with the sample and provides a means to computationally focus the fields measured in a plane remote from the object when some loss of resolution is acceptable, but a near field measurement is still required. The resultant images are free from the topographical artifacts seen in conventional near field imaging performed with shear-force feedback [25]. Furthermore, computed images allow for meaningful, quantitative interpretation of the data. The results in this Letter are encouraging and provide a proof of principle that near-field optics is amenable to computed imaging techniques as opposed to raw intensity mapping. Improved instrumentation as well as more detailed modeling are expected to provide higher resolution images with subwavelength features resolvable at distances on the order of a wavelength.

P. S. C. would like to acknowledge support under U.S. Air Force MURI Grant No. F49620-03-1-0379.

[1] D. Courjon, K. Sarayeddine, and M. Spajer, Opt. Commun. **71**, 23 (1989).
[2] C. Girard and A. Dereux, Rep. Prog. Phys. **59**, 657 (1996).
[3] J.-J. Greffet and R. Carminati, Prog. Surf. Sci. **56**, 133 (1997).
[4] E. Synge, Philos. Mag. **6**, 356 (1928).
[5] E. Ash and G. Nicholls, Nature (London) **237**, 510 (1972).
[6] A. Lewis, M. Isaacson, A. Harootunian, and A. Muray, Ultramicroscopy **13**, 227 (1984).
[7] D.W. Pohl, W. Denk, and M. Lanz, Appl. Phys. Lett. **44**, 651 (1984).
[8] E. Betzig and J. K. Trautman, Science **257**, 189 (1992).
[9] R. Dickson, D. Norris, Y.-L. Tzeng, and W. Moerner, Science **274**, 966 (1996).
[10] J.-J. Greffet, A. Sentenac, and R. Carminati, Opt. Commun. **116**, 20 (1995).
[11] N. Garcia and M. Nieto-Vesperinas, Opt. Lett. **18**, 2090 (1993).
[12] N. Garcia and M. Nieto-Vesperinas, Opt. Lett. **20**, 949 (1995).
[13] R. Carminati, J.-J. Greffet, N. Garcia, and M. Nieto-Vesperinas, Opt. Lett. **21**, 501 (1996).
[14] P. S. Carney and J. C. Schotland, Opt. Lett. **26**, 1072 (2001).
[15] P. S. Carney and J. C. Schotland, Appl. Phys. Lett. **77**, 2798 (2000).
[16] P. S. Carney and J. C. Schotland, J. Opt. A **4**, S140 (2002).
[17] S. I. Bozhevolnyi and B. Vohnsen, Phys. Rev. Lett. **77**, 3351 (1996).
[18] D. Gabor, Nature (London) **161**, 777 (1948).
[19] M. H. Maleki and A. J. Devaney, Opt. Eng. **33**, 3243 (1994).
[20] DME–DualScope, Herlev, Denmark.
[21] S. I. Bozhevolnyi, B. Vohnsen, and E. A. Bozhevolnaya, Opt. Commun. **173**, 171 (1999).
[22] B. Vohnsen and S. I. Bozhevolnyi, J. Electron Microsc. **194**, 311 (1999).
[23] S. I. Bozhevolnyi, V. A. Markel, V. Coello, W. Kim, and V. M. Shalaev, Phys. Rev. B **58**, 11 441 (1998).
[24] A. Nesci, R. Dandliker, M. Salt, and H. P. Herzig, Opt. Commun. **205**, 229 (2002).
[25] B. Hecht, H. Bielefeldt, Y. Inouye, D.W. Pohl, and L. Novotny, J. Appl. Phys. **81**, 2492 (1997).